PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C07K 14/435, G01N 33/533

A1

(11) International Publication Number: **WO 00/08054**

(43) International Publication Date: 17 February 2000 (17.02.00)

(21) International Application Number: PCT/GB99/02596

(22) International Filing Date: 6 August 1999 (06.08.99)

(30) Priority Data:

9817225.7	8 August 1998 (08.08.98)	GB
9817227.3	8 August 1998 (08.08.98)	GB
9817229.9	8 August 1998 (08.08.98)	GB

(71) Applicant *(for all designated States except US)*: IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BASTIAENS, Philippe [NL/GB]; 134a Royal College Street, Camden, London NW1 0TA (GB). PEPPERKOK, Rainer [DE/DE]; Hohl Strasse 19, D–68809 Neulussheim (DE). GELEY, Stefan [AT/GB]; 11 Frampton Road, Potters Bar, Herts EN6 1JF (GB).

(74) Agent: MILES, John; Eric Potter & Clarkson, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: MODIFIED GREEN FLUORESCENT PROTEIN

(57) Abstract

A polypeptide comprising the functional portion of green fluorescent protein but wherein valine (V) (163) has been replaced with alanine (A), serine (S) (175) has been replaced with glycine (G), isoleucine (I) (167) has been replaced with threonine (T), phenylalanine (F) (64) has been replaced with leucine (L), serine (S) (65) has been replaced with threonine (T), serine (S) (72) has been replaced with alanine (A), and threonine (T) (203) has been replaced with tyrosine (Y). Polynucleotides encoding the protein and uses of the protein as a reporter molecule are also described.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

MODIFIED GREEN FLUORESCENT PROTEIN

The present invention relates to fluorescent proteins, in particular to "green fluorescent proteins" (GFPs) and mutants thereof which have altered spectral properties compared to the wild type GFP.

Fluorescent assays for biological systems have been used for several years since, compared to many biological techniques, they have the advantage of, in some circumstances, being able to be carried out non-invasively and also they are able to give real time analysis of particular reactions in complex biological systems in which many reactions are being carried out simultaneously. In conjunction with the development of physical techniques for fluorescent assays has been the development of biological reporter constructs which serve as monitors of reactions, for example, within a cell. In particular, the development of fluorescent proteins that do not require cofactors for their intrinsic fluorescence has meant that such proteins can be introduced into and expressed in cells *via* genetic constructs.

Examples of intrinsically fluorescent proteins, notably so-called *Aequorea victoria* "green fluorescent proteins" or "GFPs" (although they may be blue or yellow) are known. Miyawaki *et al* (1997) *Nature* **388**, 882-887 describes a GFP-based Ca^{2+} sensing system; Mitra *et al* (1996) *Gene* **173**, 13-17 describes a two-GFP-based system for use in identifying protease inhibitors; WO 97/28261 discloses a two-GFP system in which the GFP donor and GFP acceptor are linked by a peptide containing a protease cleavage site. WO 95/07463 describes uses of GFPs; WO 96/23898

relates to a method of detecting biologically active substances using GFPs; Heim & Tsien (1996) *Current Biology* **6**, 178-182 relates to engineered GFPs with improved brightness, longer wavelengths and fluorescence resonance energy transfer (FRET); Poppenborg *et al* (1997) *J. Biotechnol.* **58**, 79-88 relates to GFPs as reporters for bioprocess monitoring; Park & Raines (1997) *Protein Science* **6**, 2344-2349 relates to a GFP as a signal for protein-protein interactions; Niswender *et al* (1995) *J. Microscopy* **180**, 109-116 relates to quantitative imaging of GFP in cultured cells; Chalfie *et al* (1994) *Science* **263**, 802-805 relates to GFP as a marker for gene expression; Hampton *et al* (1996) *Proc. Natl. Acad. Sci. USA* **93**, 828-833 relates to the *in vivo* examination of membrane protein localization and degradation with GFP; Heim *et al* (1995) *Nature* **373**, 663-664 relates to mutant GFPs with altered fluorescent properties; Mosser *et al* (1997) *BioTechniques* **22**, 150-161 relates to the use of a dicistronic expression cassette encoding GFP for the screening and selection of cells expressing inducible gene products; Suarez *et al* (1997) *Gene* **196**, 69-74 relates to GFP-based reporter systems for genetic analysis of bacteria; Niedenthal *et al* (1996) *Yeast* **12**, 773-778 relates to GFP as a marker for gene expression and subcellular localization in budding yeast; and Prescott *et al* (1997) *FEBS Lett* **411**, 97-101 relates to the use of GFP as a marker for assembled mitochondrial ATP synthase in yeast. GFPs and their uses have been reviewed in Pozzan *et al* (1997) *Nature* **388**, 834-835, Misteli & Spector (1997) *Nature Biotechnology* **15**, 961-964; and Cubitt *et al* (1995) *Trends Biochem. Sci.* **20**, 448-455.

Despite many variant GFPs being known, there continues to be the need for variant GFPs which have improved properties, especially improved or

new spectral properties, for use in biological systems, especially those where fluorescence resonance energy transfer (FRET) is used to study the biological system.

An example of a FRET-based method for studying biological systems is described in detail in UK Patent Application No 9817229.9 entitled "Fluorescent assay for biological systems" and the PCT application claiming priority from that application and which was filed on the same day as this application. A further example of a FRET-based method is described in Miyawaki *et al* (1997) *Nature* **388**, 882-887.

A first aspect of the invention provides a polypeptide comprising the functional portion of green fluorescent protein but wherein valine (V) 163 has been replaced with alanine (A), serine (S) 175 has been replaced with glycine (G), isoleucine (I) 167 has been replaced with threonine (T), phenylalanine (F) 64 has been replaced with leucine (L), serine (S) 65 has been replaced with threonine (T), serine (S) 72 has been replaced with alanine (A), and threonine (T) 203 has been replaced with tyrosine (Y).

By "functional portion" we include the meaning that it is the portion of the protein which, in the absence of other portions of GFP, gives rise to useful fluorescent properties, such as the portion being fluorescent. It will be appreciated that, in respect of this first aspect of the invention the GFP or polypeptide comprising the functional portion of GFP with the given mutations may also include other mutations which may confer further desirable properties.

A second aspect of the invention provides a polypeptide which has the amino acid sequence

MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTY
GVQCFARYPDHMKR
HDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN
SHNVYIMADKQKNG
IKANFKTRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSQDPHGKRDHMVLLEFVTS
AGITHGMDELYKSF

This is the sequence shown in Figure 3.

Conveniently, the difference with respect to the native GFP can be summarised as S2G, H25Q, S30R, F64L, S65T, S72A, Q80R, F84L, V163A, I167T, S175G, T203Y, K209Q, N212H, E213G, A216S, + 239S + 240F, making use of standard single letter amino acid code.

By "green fluorescent protein", in the context of a single protein, we mean wild-type green fluorescent protein as described in Prasher *et al* (1992) *Gene* **111**, 229-233 and whose amino acid sequence is given in Figure 1. As noted above, the term GFP may be used to denote variants which in fact fluoresce yellow or blue.

The sequence of a particular preferred polypeptide of the invention (variant GFP) is shown in Figure 3, and its sequence is compared with *A. victoria* GFP and mm GFP5 (Zernicka-Goeta *et al* (1997) *Development* **124**, 1133-1137 in Figures 4 and 5.

It will be appreciated that the functional portion of the polypeptide which contains the mutations as said may be incorporated into any suitable

polypeptide in which it is desired to have a fluorescent moiety. Typically, the functional portion is included in a polypeptide whose fluorescence or change in fluorescence is measured under suitable conditions. Thus, the polypeptide may be one which is expressed as a reporter molecule (since its expression may be measured fluorimetrically). Alternatively, the functional portion may be included in a polypeptide which is used in a biological system which makes use of FRET. For example, a polypeptide of the invention includes a polypeptide which contains the fluorescent portion as said, and it is used in conjunction with another fluorescent moiety with which it acts as a donor or acceptor in a FRET reaction. Most suitably, the polypeptide of the invention contains, in addition to the functional portion as said, a further fluorescent moiety in the same polypeptide chain and the pair of fluorescent moieties may act as donor-acceptor pairs in a FRET reaction. Thus, the polypeptide of the invention typically is a fusion protein containing at least the functional portion of the polypeptide which contains the mutations as said.

Thus, the polypeptide of the invention may be used in any suitable prior art FRET method.

The polypeptides of the invention are particularly suited for use in the FRET method described in UK Patent Application No 9817229.9 entitled "Fluorescent assay for biological systems" and the PCT application claiming priority from that application and which was filed on the same day as this application since the mutations confer an unusually high fluorescent lifetime. The polypeptides of the invention are believed to be particularly suited as acceptor molecules since, at least in relation to the

molecule of the second aspect of the invention, it excites at 514nm and emits at 531nm.

A particularly preferred polypeptide of the invention is one which has the amino acid sequence as shown in Figure 3.

A further preferred polypeptide of the invention is one which comprises at least residues 7 to 229 of green fluorescent protein containing said amino acid replacements. The minimal domain required for fluorescence in GFP is believed to be amino acids 7 to 229 (Li *et al* (1997) *J. Biol. Chem.* **272**, 28545-28549. Also, this information, and other information about GFPs, is available from Clontech Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA (xqli@CLONTECH.com).

It is preferred that a polypeptide which comprises at least residues 7 to 229 of green fluorescent protein not only contains the amino acid replacements described with respect to the first aspect of the invention, but also, compared to native *A. victoria* GFP, the amino acid replacements H25Q, S30R, Q80R, F84L, K209Q, N212H, F213G and A216S.

A still further preferred polypeptide is one comprising a further fluorescent moiety. In particular, the further fluorescent moiety is one which is capable of FRET with the said portion of the variant GFP. Typically, this further fluorescent moiety is a GFP or a variant GFP. Thus, a particularly preferred polypeptide of the invention is a fusion polypeptide which contains at least the functional portion of the

polypeptide which contains the mutations as said, and contains a further mutant GFP.

A third aspect of the invention provides a polynucleotide encoding a polypeptide of the first or second aspect of the invention. The polynucleotide may be DNA or RNA; DNA is preferred. A particularly preferred polynucleotide of the invention is shown in Figure 3 (DNA sequence) but, because of the degeneracy of the genetic code, it will be appreciated that other polynucleotides may encode the same polypeptide (ie with the amino acid sequence given in Figure 3).

A fourth aspect of the invention provides an expression vector encoding a polypeptide of the first or second aspect of the invention.

The expression vectors of the invention, and other polynucleotides can be constructed by standard laboratory molecular biology methods such as those described in Sambrook *et al* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York incorporated herein by reference.

The polynucleotide of the invention (typically DNA) may be expressed in a suitable host to produce a polypeptide comprising the polypeptide of the invention. Thus, the DNA encoding the polypeptide of the invention may be used in accordance with known techniques, appropriately modified in view of the teachings contained herein, to construct an expression vector, which is then used to transform an appropriate host cell for the expression and production of the polypeptide of the invention. Such techniques include

those disclosed in US Patent Nos. 4,440,859 issued 3 April 1984 to Rutter *et al*, 4,530,901 issued 23 July 1985 to Weissman, 4,582,800 issued 15 April 1986 to Crowl, 4,677,063 issued 30 June 1987 to Mark *et al*, 4,678,751 issued 7 July 1987 to Goeddel, 4,704,362 issued 3 November 1987 to Itakura *et al*, 4,710,463 issued 1 December 1987 to Murray, 4,757,006 issued 12 July 1988 to Toole, Jr. *et al*, 4,766,075 issued 23 August 1988 to Goeddel *et al* and 4,810,648 issued 7 March 1989 to Stalker, all of which are incorporated herein by reference.

The polynucleotide, such as DNA, encoding the polypeptide of the invention may be joined to a wide variety of other DNA sequences for introduction into an appropriate host. The companion DNA will depend upon the nature of the host, the manner of the introduction of the DNA into the host, and whether episomal maintenance or integration is desired.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through standard techniques. Generally, not all of the hosts will be transformed by the vector. Therefore, it will be necessary to select for transformed host cells. One selection technique involves incorporating into the expression vector a DNA sequence, with any necessary control elements, that codes for a selectable trait in the transformed cell, such as antibiotic resistance.

Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

Host cells that have been transformed by the recombinant DNA of the invention are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the expression of the polypeptide, which can then be recovered.

Many expression systems are known, including bacteria (for example *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*), filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

The vectors include a prokaryotic replicon, such as the ColE1 *ori*, for propagation in a prokaryote, even if the vector is to be used for expression in other, non-prokaryotic, cell types. The vectors can also include an appropriate promoter such as a prokaryotic promoter capable of directing the expression (transcription and translation) of the genes in a bacterial host cell, such as *E. coli*, transformed therewith.

A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with exemplary bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention.

Typical prokaryotic vector plasmids are pUC18, pUC19, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA, USA) and p*Trc*99A and pKK223-3 available from Pharmacia, Piscataway, NJ, USA.

A typical mammalian cell vector plasmid is pSVL available from Pharmacia, Piscataway, NJ, USA. This vector uses the SV40 late promoter to drive expression of cloned genes, the highest level of expression being found in T antigen-producing cells, such as COS-1 cells.

An example of an inducible mammalian expression vector is pMSG, also available from Pharmacia. This vector uses the glucocorticoid-inducible promoter of the mouse mammary tumour virus long terminal repeat to drive expression of the cloned gene.

Useful yeast plasmid vectors are pRS403-406 and pRS413-416 and are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Plasmids pRS403, pRS404, pRS405 and pRS406 are Yeast Integrating plasmids (YIps) and incorporate the yeast selectable markers *HIS3, TRP1, LEU2* and *URA3*. Plasmids pRS413-416 are Yeast Centromere plasmids (YCps).

A variety of methods have been developed to operably link DNA to vectors via complementary cohesive termini. For instance, complementary homopolymer tracts can be added to the DNA segment to be inserted into the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, generated by endonuclease restriction digestion, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerizing activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a large molar excess of linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including International Biotechnologies Inc, New Haven, CN, USA.

A desirable way to modify the DNA encoding the polypeptide of the invention is to use the polymerase chain reaction as disclosed by Saiki *et al* (1988) *Science* **239**, 487-491.

In this method the DNA to be enzymatically amplified is flanked by two specific oligonucleotide primers which themselves become incorporated into the amplified DNA. The said specific primers may contain restriction endonuclease recognition sites which can be used for cloning into expression vectors using methods known in the art.

The present invention also relates to a host cell transformed with a polynucleotide vector construct of the present invention. The host cell can be either prokaryotic or eukaryotic. Bacterial cells are preferred prokaryotic host cells and typically are a strain of *E. coli* such as, for example, the *E. coli* strains DH5 available from Bethesda Research Laboratories Inc., Bethesda, MD, USA, and RR1 available from the American Type Culture Collection (ATCC) of Rockville, MD, USA (No ATCC 31343). Preferred eukaryotic host cells include yeast and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human fibroblastic cell line. Yeast host cells include YPH499, YPH500 and YPH501 which are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Preferred mammalian host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CCL61, NIH Swiss mouse embryo cells NIH/3T3 available from the ATCC as CRL 1658, and monkey kidney-derived COS-1 cells available from the ATCC as CRL 1650.

Transformation of appropriate cell hosts with a DNA construct of the present invention is accomplished by well known methods that typically depend on the type of vector used. With regard to transformation of prokaryotic host cells, see, for example, Cohen *et al* (1972) *Proc. Natl.*

Acad. Sci. USA **69**, 2110 and Sambrook *et al* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. Transformation of yeast cells is described in Sherman *et al* (1986) *Methods In Yeast Genetics, A Laboratory Manual*, Cold Spring Harbor, NY. The method of Beggs (1978) *Nature* **275**, 104-109 is also useful. With regard to vertebrate cells, reagents useful in transfecting such cells, for example calcium phosphate and DEAE-dextran or liposome formulations, are available from Stratagene Cloning Systems, or Life Technologies Inc., Gaithersburg, MD 20877, USA.

Electroporation is also useful for transforming cells and is well known in the art for transforming yeast cells, bacterial cells and vertebrate cells.

For example, many bacterial species may be transformed by the methods described in Luchansky *et al* (1988) *Mol. Microbiol.* **2**, 637-646 incorporated herein by reference. The greatest number of transformants is consistently recovered following electroporation of the DNA-cell mixture suspended in 2.5X PEB using 6250V per cm at 25μFD.

Methods for transformation of yeast by electroporation are disclosed in Becker & Guarente (1990) *Methods Enzymol.* **194**, 182.

Successfully transformed cells, ie cells that contain a DNA construct of the present invention, can be identified by well known techniques. For example, cells resulting from the introduction of an expression construct of the present invention can be grown to produce the polypeptide of the invention. Cells can be harvested and lysed and their DNA content

examined for the presence of the DNA using a method such as that described by Southern (1975) *J. Mol. Biol.* **98**, 503 or Berent *et al* (1985) *Biotech.* **3**, 208. Alternatively, the presence of the protein in the supernatant can be detected using antibodies.

However, a convenient way of identifying transformed cells which express the polypeptide is that they are fluorescent.

In addition to directly assaying for the presence of recombinant DNA, successful transformation can be confirmed by well known immunological methods when the recombinant DNA is capable of directing the expression of the protein. For example, cells successfully transformed with an expression vector produce proteins displaying appropriate antigenicity. Samples of cells suspected of being transformed are harvested and assayed for the protein using suitable antibodies. Of course, transformation and expression is indicated by the production of a fluorescent protein in this case.

Thus, in addition to the transformed host cells themselves, the present invention also contemplates a culture of those cells, preferably a monoclonal (clonally homogeneous) culture, or a culture derived from a monoclonal culture, in a nutrient medium.

A particularly suitable "starting" vector is the pcDNA3.1 vector distributed by Invitrogen (Invitrogen BV, De Schelp 12, 9351 NV Leek, The Netherlands). The key features of this vector for this invention are: (i) Cytomegalovirus enhancer-promoter for high level expression of the insert

in mammalian cells (the insert is for example the cDNA encoding the polypeptides described above and has to be cloned into the vector); (ii) multiple cloning site in forward and reverse orientation; (iii) expression cassette for a selectable marker in eukaryotic cells (neomycin, zeocin or hygromycin).

When the polypeptide of the invention is to be used *ex vivo*, such as in an *in vitro* assay or the like, it may be convenient to express the polypeptide in a bacterial system (such as *E. coli*), or in yeast or insect cells, or in other systems which have been designed for facile production of large amounts of protein. When the polypeptide of the invention is to be used in an *in vivo* assay it is conveniently expressed in the cell in which the assay is to be carried out, typically a mammalian cell.

In a particularly preferred embodiment, DNA encoding the polypeptide of the invention (such as YFP5) may be fused to the promoter/enhancer elements of a gene under investigation. Such DNA stably introduced into mammalian cells may be used as a reporter for expression of the respective gene under investigation. Readout of the activity is the amount of polypeptide, such as YFP5, expressed, which can be determined by determination of the specific fluorescence of the polypeptide. Similar DNAs may be generated for the other GFPs such as those listed in Table 1 (see below). Since they have overlapping spectra they cannot be used simultaneously in the same cells. However, using fluorescent lifetime imaging CFP, MmGFP5 and YFP5, for example, could be used simultaneously as their lifetimes are sufficiently separated from each other. Using multiple frequency FLIM (fluorescent lifetime imaging) the

relative amounts of these three GFP mutants expressed in the same cell could be determined with high precision and hence the promoter activity of at least three genes.

Multifrequency FLIM is described in UK Patent Application No 9817227.3 entitled "Multiple Frequency Fluorescence Lifetime Imaging" and the PCT application claiming priority from that application and which has the same filing date as this application.

The invention will now be described in more detail with reference to the following Figures and Example wherein:

Figure 1 shows the cDNA and amino acid sequence of *A. victoria* green fluorescent protein (GFP).

Figure 2 shows the cDNA and amino acid sequence of a prior art mutant GFP (mmGFP5; Zernicka-Goetz *et al*).

Figure 3 shows the cDNA and amino acid sequence of a polypeptide of the invention (called mmYFP or mYFP5 or YFP5) which is described in more detail in Example 1.

Figure 4 is a comparison of the cDNA sequences from Figures 1 to 3.

Figure 5 is a comparison of the amino acid sequences from Figures 1 to 3.

Example 1: Construction of mutant GFP and its properties

The mutant GFP which we call YFP5, which is a red-shifted mutant of MmGFP5, was generated by PCR-mediated site-directed mutagenesis of MmGFP5 (Zernicka-Goetz *et al* (1997) *Development* **124**, 1133-1137). MmGFP5 is a wtGFP mutated in V163A, S175G, I167T, F64L and S65T; the mutations V163A, S175G and I167T were introduced into wtGFP by Siemering *et al* (1996) *Current Biol.* **6**, 1653-, and Zernicka-Goetz *et al* introduced the mutations F64L and S65T). This approach introduced mutations S72A and T203Y into MmGFP5 using primer pairs ATGCGGCCGCGAATTCGCCACCATGGGTAAAGGAGAAGAACTT and CTGGGTATCTTGCGAAGCATTGTACGTACAATGCTTCGCAAGATACCCAG; and GAAAGGGCAGATTGATAGGACAGGTAATGCATTACCTGTCCTATAATCTGCCCTT TC and AAGGATCCTCTAGAAGCTTTTGTATAGTTCATCCATG. The underlined nucleotides indicate mismatches.

The fluorescent lifetimes of various GFP mutants are shown in Table 1.

References to Table 1

1. Heim & Tsien (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Bio.* **6**, 178-182.

2. Orme M *et al* (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392-1395.

3. Zernicka-Goetz *et al* (1997). Following cell fate in the living mouse embryo. *Development* **124**, 1133-1137.

4. Miyawaki *et al* (1997). Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**, 882-887.

The final PCR product was gel-purified, digested with *Eco*RI and *Xba*I and subcloned into pEFT7MCS. This vector is based on pEF-BOS (*Nucleic Acids Res.* (1990) Sep 11; **18(17)**, 5322 pEF-BOS, a powerful mammalian expression vector. Mizushima S, Nagata S). A modified version of pEF-BOS containing a Neo resistance expression cassette, pEF1-Neo, was obtained from G. Baier, Innsbruck. The Neo expression cassette to make the vector smaller and introduced a T7 RNA polymerase promoter as well as several unique restriction enzyme sites downstream of the human EF1α promoter and the SV40 polyadenylation site.

Any other suitable vector, as described in the specification, may be used for the expression of the mutant GFP. The introduced mutations were verified by sequencing using Sequenase. The sequence of YFP5 is given in Figure 3.

The respective GFP mutants were expressed in cells by microinjection (Pepperkok *et al*, 1997 in "Microinjection and Transgenesis", eds. Cid-Arregui and Garcia-Carranca, Springer, Heidelberg, pp 145-154) of plasmids based on the vector pEFT7MCS and with inserts of the respective GFP encoding cDNAs. At 2h after microinjection cells were mounted on the FLIM microscope set-up and the respective lifetimes were

determined at 37°C in living cells. Any suitable expression system or lifetime-detection system may be used.

YFP5 shows a well-separated and significantly longer lifetime than other GFP mutants making it an ideal partner in multi-labelling FLIM experiments.

Table 1: Fluorescent lifetimes of various GFP mutants.

Name of GFP	Exitation peak (nm)	Emission peak (nm)	Fluorescent lifetime (ns) tf / tm	Reference/ source	Mutations
S65T	489	511	2.57/2.59	Heim and Tsien	S65T
YFP-10C	513	527	2.85/2.88	Orme et al./ Clontech	S65G, V68L, S72A, T203Y
MmGFP5	473	507	2.42/2.68	Zernicka-Goetz, et al.	F64L, S65T,V163A, I167T, S175G,
CFP*	432 (453)	476 (503)	1.32/2.23	Miyawaki et al.	F64L, S65T, Y66W, N146I, M153T, V163A, N212K
YFP5	514	531	3.69/3.60	This work	F64L, S65T, S72A, V163A, I167T, S175G, T203Y

* : numbers in parenthesis are the side-peaks in excitation and emission of CFP which are used in the single excitation wavelength method to measure FRET by "ingrowth". The fluorescent life-time was measured by excitation at 488nm.

CLAIMS

1. A polypeptide comprising the functional portion of green fluorescent protein but wherein valine (V) 163 has been replaced with alanine (A), serine (S) 175 has been replaced with glycine (G), isoleucine (I) 167 has been replaced with threonine (T), phenylalanine (F) 64 has been replaced with leucine (L), serine (S) 65 has been replaced with threonine (T), serine (S) 72 has been replaced with alanine (A), and threonine (T) 203 has been replaced with tyrosine (Y).

2. A polypeptide which has the amino acid sequence

MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTY
GVQCFARYPDHMKR
HDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN
SHNVYIMADKQKNG
IKANFKTRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSQDPHGKRDHMVLLEFVTS
AGITHGMDELYKSF.

3. A polypeptide according to Claim 1 which has the additional amino acid replacements S2G, H25Q, S30R, Q80R, F84L, N212H, E213G, A216S, and additional residues 239S and 240F.

4. A polypeptide comprising at least residues 7 to 229 of green fluorescent protein containing the amino acid replacements as defined in any of Claims 1 to 3.

5. A polypeptide according to any one of Claims 1, 3 and 4 comprising a further fluorescent moiety.

6. A polypeptide according to Claim 5 wherein the further fluorescent moiety is a green fluorescent protein or a variant thereof.

7. A polypeptide according to any one of Claims 1, and 3 to 6 which is a fusion polypeptide.

8. A polypeptide according to Claim 7 wherein the fusion polypeptide is one used in a biological system which makes use of FRET.

9. A polynucleotide encoding a polypeptide according to any one of the preceding claims.

10. An expression vector encoding a polypeptide according to any one of Claims 1 to 8.

11. A host cell comprising a polynucleotide according to Claim 9 or an expression vector according to Claim 10.

12. Use of a polypeptide according to any one of Claims 1 to 8 as a reporter molecule in a cell.

13. Use of a polynucleotide according to Claim 9 or an expression vector according to Claim 8 to express a reporter molecule in a cell.

14. Any novel fluorescent protein as herein described.

Figure 1

A.victoria GFP:

cDNA:

atgagtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaatt
ttctgtcagtggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactggaaaactac
ctgttccatggccaacacttgtcactactttctcttatggtgttcaatgcttttcaagatacccagatcatatgaaacag
catgactttttcaagagtgccatgcccgaaggttatgtacaggaaagaactatattttcaaagatgacgggaactacaa
gacacgtgctgaagtcaagtttgaaggtgataccttgttaatagaatcgagttaaaaggtattgattttaaagaagatg
gaaacattcttggacacaaattggaatacaactataactcacacaatgtatacatcatggcagacaaacaaaagaatgga
atcaaagttaacttcaaaattagacacaacattgaagatggaagcgttcaactagcagaccattatcaacaaaatactcc
aattggcgatggccctgtccttttaccagacaaccattacctgtccacacaatctgccctttcgaaagatcccaacgaaa
agagagaccacatggtccttcttgagtttgtaacagctgctgggattacacatggcatggatgaactatacaaataa

protein:

MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQ
HDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG
IKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYKZ

Figure 2

mmGFP5:

CDNA:

atgggtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcaaaaatt
ctctgtcaggggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactgggaagctac
ctgttccctggccaacacttgtcactactttgacttatggtgtacaatgcttctcaagatacccagatcatatgaagcgg
cacgacttcctcaagagcgccatgcctgagggatacgtgcaggagaggaccatcttcttcaaggacgacgggaactacaa
gacacgtgctgaagtcaagtttgagggagacaccctcgtcaacaggatcgagcttaagggaatcgatttcaaggaggacg
gaaacatcctcggccacaagttggaatacaactacaactcccacaacgtatacatcatggccgacaagcaaaagaacggc
atcaaagccaacttcaagacccgccacaacatcgaagacggcggcgtgcaactcgctgatcattatcaacaaaatactcc
aattggcgatggccctgtccttttaccagacaaccattacctgtccacacaatctgccctttcccaagatccccacggaa
agagagatcacatggtccttcttgagtttgttacatctgctgggattacacatggcatggatgaactatacaaaagcttc
tag

protein:

MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKR
HDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG
IKANFKTRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYKSF
Z

Figure 3

mmYFP:

cDNA:

atgggtaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcaaaaatt
ctctgtcaggggagagggtgaaggtgatgcaacatacggaaaacttacccttaaatttatttgcactactgggaagctac
ctgttccctggccaacacttgtcactactttgacttatggtgtacaatgcttcGcaagatacccagatcatatgaagcgg
cacgacttcctcaagagcgccatgcctgagggatacgtgcaggagaggaccatcttcttcaaggacgacgggaactacaa
gacacgtgctgaagtcaagtttgagggagacaccctcgtcaacaggatcgagcttaagggaatcgatttcaaggaggacg
gaaacatcctcggccacaagttggaatacaactacaactcccacaacgtatacatcatggccgacaagcaaaagaacggc
atcaaagccaacttcaagacccgccacaacatcgaagacggcggcgtgcaactcgctgatcattatcaacaaaatactcc
aattggcgatggccctgtccttttaccagacaaccattacctgtccTATcaatctgccctttcccaagatccccacggaa
agagagatcacatggtccttcttgagtttgttacatctgctgggattacacatggcatggatgaactatacaaaagcttc
tag

protein:

MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFARYPDHMKR
HDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNG
IKANFKTRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYKSF
Z

Figure 4 (page 1 of 3)

Nucleic acid alignment:

```
                    . . . .10 . . . .20 . . . .30 . . . .40 . . . .50 . . . .60
A.vict.GFP      1:............ATGaGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTT: 48
MMGFP5          1:GAATTCGCCACCATGGGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTT: 60
MMYFP5          1:GAATTCGCCACCATGGGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTT: 60
consensus       1:************!!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!: 60

                    . . . .70 . . . .80 . . . .90 . . . 100 . . . 110 . . . 120
A.vict.GFP     49:GAATTAGATGGTGATGTTAATGGGCAcAAATTtTCTGTCAGtGGAGAGGGTGAAGGTGAT:108
MMGFP5         61:GAATTAGATGGTGATGTTAATGGGCAAAAATTCTCTGTCAGGGGAGAGGGTGAAGGTGAT:120
MMYFP5         61:GAATTAGATGGTGATGTTAATGGGCAAAAATTCTCTGTCAGGGGAGAGGGTGAAGGTGAT:120
consensus      61:!!!!!!!!!!!!!!!!!!!!!!!!!!!!!*!!!!!*!!!!!!!!*!!!!!!!!!!!!!!!:120

                    . . . 130 . . . 140 . . . 150 . . . 160 . . . 170 . . . 180
A.vict.GFP    109:GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGaAAaCTACCTGTTCCa:168
MMGFP5        121:GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGGAAGCTACCTGTTCCC:180
MMYFP5        121:GCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGGAAGCTACCTGTTCCC:180
consensus     121:!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!*!!*!!!!!!!!!!!*:180

                    . . . 190 . . . 200 . . . 210 . . . 220 . . . 230 . . . 240
A.vict.GFP    169:TGGCCAACACTTGTCACTACTTTctCTTATGGTGTtCAATGCTTtTCAAGATACCCAGAT:228
MMGFP5        181:TGGCCAACACTTGTCACTACTTTGACTTATGGTGTACAATGCTTCTCAAGATACCCAGAT:240
MMYFP5        181:TGGCCAACACTTGTCACTACTTTGACTTATGGTGTACAATGCTTCgCAAGATACCCAGAT:240
consensus     181:!!!!!!!!!!!!!!!!!!!!!!!!!**!!!!!!!!!!*!!!!!!!!**!!!!!!!!!!!!!:240

                    . . . 250 . . . 260 . . . 270 . . . 280 . . . 290 . . . 300
A.vict.GFP    229:CATATGAAaCaGCAtGACTTttTCAAGAGtGCCATGCCcGAaGGtTAtGTaCAGGAaAGa:288
MMGFP5        241:CATATGAAGCGGCACGACTTCCTCAAGAGCGCCATGCCTGAGGGATACGTGCAGGAGAGG:300
MMYFP5        241:CATATGAAGCGGCACGACTTCCTCAAGAGCGCCATGCCTGAGGGATACGTGCAGGAGAGG:300
consensus     241:!!!!!!!!!*!*!!!*!!!!!**!!!!!!!*!!!!!!!!*!!*!!*!!*!!*!!!!!*!!*:300

                    . . . 310 . . . 320 . . . 330 . . . 340 . . . 350 . . . 360
```

Figure 4 (page 2 of 3)

```
A.vict.GFP 289:ACtATaTTtTTCAAaGAtGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAaGGt:348
MMGFP5     301:ACCATCTTCTTCAAGGACGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAGGGA:360
MMYFP5     301:ACCATCTTCTTCAAGGACGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAGGGA:360
consensus  301:!!*!!*!!*!!!!!*!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!*!!*:360

                    . . . 370 . . . 380 . . . 390 . . . 400 . . . 410 . . . 420
A.vict.GFP 349:GAtACCCTtGTtAAtAGaATCGAGtTaAAaGGtATtGATTTtAAaGAaGAtGGAAACATt:408
MMGFP5     361:GACACCCTCGTCAACAGGATCGAGCTTAAGGGAATCGATTTCAAGGAGGACGGAAACATC:420
MMYFP5     361:GACACCCTCGTCAACAGGATCGAGCTTAAGGGAATCGATTTCAAGGAGGACGGAAACATC:420
consensus  361:!!*!!!!!*!!*!!*!!*!!!!!!*!*!!*!!*!!*!!!!!*!!*!!*!!*!!!!!!!!*:420

                    . . . 430 . . . 440 . . . 450 . . . 460 . . . 470 . . . 480
A.vict.GFP 409:CTtGGaCACAAaTTGGAATACAACTAtAACTCaCACAAtGTATACATCATGGCaGACAAa:468
MMGFP5     421:CTCGGCCACAAGTTGGAATACAACTACAACTCCCACAACGTATACATCATGGCCGACAAG:480
MMYFP5     421:CTCGGCCACAAGTTGGAATACAACTACAACTCCCACAACGTATACATCATGGCCGACAAG:480
consensus  421:!!*!!*!!!!!*!!!!!!!!!!!!!!!*!!!!!*!!!!!*!!!!!!!!!!!!!!!*!!!!!*:480

                    . . . 490 . . . 500 . . . 510 . . . 520 . . . 530 . . . 540
A.vict.GFP 469:CAAAAGAAtGGaATCAAAGttAACTTCAAaAttaGaCACAACATtGAAGAtGGaaGCGTt:528
MMGFP5     481:CAAAAGAACGGCATCAAAGCCAACTTCAAGACCCGCCACAACATCGAAGACGGCGGCGTG:540
MMYFP5     481:CAAAAGAACGGCATCAAAGCCAACTTCAAGACCCGCCACAACATCGAAGACGGCGGCGTG:540
consensus  481:!!!!!!!!!*!!*!!!!!!!**!!!!!!!!!*!***!*!!!!!!!!!*!!!!!*!!**!!!!*:540

                    . . . 550 . . . 560 . . . 570 . . . 580 . . . 590 . . . 600
A.vict.GFP 529:CAACTaGCaGAcCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA:588
MMGFP5     541:CAACTCGCTGATCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA:600
MMYFP5     541:CAACTCGCTGATCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCA:600
consensus  541:!!!!!*!!*!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!:600

                    . . . 610 . . . 620 . . . 630 . . . 640 . . . 650 . . . 660
A.vict.GFP 589:GACAACCATTACCTGTCCACACAATCTGCCCTTTCgaAAGATCCCaACGaAAAGAGAGAc:648
MMGFP5     601:GACAACCATTACCTGTCCACACAATCTGCCCTTTCCCAAGATCCCCACGGAAAGAGAGAT:660
MMYFP5     601:GACAACCATTACCTGTCCtatCAATCTGCCCTTTCCCAAGATCCCCACGGAAAGAGAGAT:660
consensus  601:!!!!!!!!!!!!!!!!!!!!***!!!!!!!!!!!!!!!!!*!!!!!!!!*!!!*!!!!!!!!*:660
```

Figure 4 (page 3 of 3)

```
                  . . . 670 . . . 680 . . . 690 . . . 700 . . . 710 . . . 720
A.vict.GFP 649:CACATGGTCCTTCTTGAGTTTGTaACAgCTGCTGGGATTACACATGGCATGGATGAACTA:708
MMGFP5     661:CACATGGTCCTTCTTGAGTTTGTTACATCTGCTGGGATTACACATGGCATGGATGAACTA:720
MMYFP5     661:CACATGGTCCTTCTTGAGTTTGTTACATCTGCTGGGATTACACATGGCATGGATGAACTA:720
consensus  661:!!!!!!!!!!!!!!!!!!!!!!!!!*!!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!!:720


                  . . . 730 . . .
A.vict.GFP 709:TACAAAtaa.......:717
MMGFP5     721:TACAAAAGCTTCTAGA:736
MMYFP5     721:TACAAAAGCTTCTAGA:736
consensus  721:!!!!!!**********:736
```

Figure 5

protein alignment:

```
A.vict.      1:MsKGEELFTGVVPILVELDGDVNGhKFSVsGEGEGDATYGKLTLKFICTTGKLPVPWPTL: 60
mmGFP5       1:MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTL: 60
mmYFP5       1:MGKGEELFTGVVPILVELDGDVNGQKFSVRGEGEGDATYGKLTLKFICTTGKLPVPWPTL: 60
consensus    1:!*!!!!!!!!!!!!!!!!!!!!!!!*!!!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!: 60

                  . . . .70 . . . .80 . . . .90 . . . 100 . . . 110 . . . 120
A.vict.     61:VTTfsYGVQCFSRYPDHMKqHDFfKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV:120
mmGFP5      61:VTTLTYGVQCFSRYPDHMKRHDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV:120
mmYFP5      61:VTTLTYGVQCFaRYPDHMKRHDFLKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV:120
consensus   61:!!!**!!!!!!*!!!!!!!*!!!*!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!:120

                  . . . 130 . . . 140 . . . 150 . . . 160 . . . 170 . . . 180
A.vict.    121:NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKvNFKiRHNIEDGsVQLAD:180
mmGFP5     121:NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKTRHNIEDGGVQLAD:180
mmYFP5     121:NRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANFKTRHNIEDGGVQLAD:180
consensus  121:!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!*!!!*!!!!!!!*!!!!!:180

                  . . . 190 . . . 200 . . . 210 . . . 220 . . . 230 . . . 240
A.vict.    181:HYQQNTPIGDGPVLLPDNHYLSTQSALSkDPneKRDHMVLLEFVTaAGITHGMDELYKz.:238
mmGFP5     181:HYQQNTPIGDGPVLLPDNHYLSTQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYKSF:240
mmYFP5     181:HYQQNTPIGDGPVLLPDNHYLSyQSALSQDPHGKRDHMVLLEFVTSAGITHGMDELYKSF:240
consensus  181:!!!!!!!!!!!!!!!!!!!!!!!*!!!!!*!!**!!!!!!!!!!!!!*!!!!!!!!!!!**:240
```

SEQUENCE LISTING

<110> Bastiaens, Philippe
Pepperkok, Rainer
Geley, Stefan
Imperial College Research Technology Limited

<120> Fluorescent protein

<130> IMPWP21223

<140>
<141>

<150> GB 9817225.7
<151> 1998-08-08

<160> 6

<170> PatentIn Ver. 2.0

<210> 1
<211> 722
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Mutant GFP

<400> 1
atgggtaaag gagaagaact tttcactgga gttgtcccaa ttcttgttga attagatggt 60
gatgttaatg ggcaaaaatt ctctgtcagg ggagagggtg aaggtgatgc aacatacgga 120
aaacttaccc ttaaatttat ttgcactact gggaagctac tgttccctgg ccaacacttg 180
tcactacttt gacttatggt gtacaatgct tcgcaagata cccagatcat atgaagcggc 240
acgacttcct caagagcgcc atgcctgagg gatacgtgca ggagaggacc atcttcttca 300
aggacgacgg gaactacaag acacgtgctg aagtcaagtt tgagggagac accctcgtca 360
acaggatcga gcttaaggga atcgatttca aggaggacgg aaacatcctc ggccacaagt 420
tggaatacaa ctacaactcc cacaacgtat acatcatggc cgacaagcaa aagaacggca 480
tcaaagccaa cttcaagacc cgccacaaca tcgaagacgg cggcgtgcaa ctcgctgatc 540
attatcaaca aaatactcca attggcgatg gccctgtcct tttaccagac aaccattacc 600
tgtcctatca atctgccctt tcccaagatc cccacggaaa gagagatcac atggtccttc 660
ttgagtttgt tacatctgct gggattacac atggcatgga tgaactatac aaaagcttct 720
ag 722

<210> 2
<211> 240
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Mutant GFP

<400> 2

```
Met Gly Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val
  1               5                  10                  15

Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Arg Gly Glu
             20                  25                  30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
         35                  40                  45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu
     50                  55                  60

Thr Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys Arg
 65                  70                  75                  80

His Asp Phe Leu Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                 85                  90                  95

Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
            100                 105                 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
        115                 120                 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
    130                 135                 140

Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
145                 150                 155                 160

Ile Lys Ala Asn Phe Lys Thr Arg His Asn Ile Glu Asp Gly Gly Val
                165                 170                 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
            180                 185                 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu Ser
        195                 200                 205

Gln Asp Pro His Gly Lys Arg Asp His Met Val Leu Leu Glu Phe Val
    210                 215                 220

Thr Ser Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys Ser Phe
```

225 230 235 240

<210> 3
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR primer

<400> 3
atgcggccgc gaattcgcca ccatgggtaa aggagaagaa ctt 43

<210> 4
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR primer

<400> 4
ctgggtatct tgcgaagcat tgtacgtaca atgcttcgca agatacccag 50

<210> 5
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR primer

<400> 5
gaaagggcag attgatagga caggtaatgc attacctgtc ctataatctg ccctttc 57

<210> 6
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:PCR primer

<400> 6

aaaaggatcc tctagaagct tttgtatagt tcatccatg 39

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 99/02596

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/435 G01N33/533

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 96 27675 A (MEDICAL RES COUNCIL ;HASELOFF JAMES PHILLIP (AU); HODGE SARAH (GB)) 12 September 1996 (1996-09-12) *Seq. ID. 10* page 31, line 11; claims 11-15	
A	WO 98 06737 A (HEIM ROGER ;CUBITT ANDREW B (GB); ORMO MATS F (SE); REMINGTON JAME) 19 February 1998 (1998-02-19) claims	
A	WO 97 11094 A (NOVONORDISK AS ;THASTRUP OLE (DK); TULLIN SOEREN (DK); POULSEN LAR) 27 March 1997 (1997-03-27) page 3, line 13	

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 November 1999	30/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cervigni, S

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 99/02596

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passagee	Relevant to claim No.
A	US 5 625 048 A (TSIEN ROGER Y ET AL) 29 April 1997 (1997-04-29) claims 8,12	
A	WO 98 30715 A (ISACOFF EHUD Y ;SIEGAL MICAH S (US); UNIV CALIFORNIA (US); CALIFOR) 16 July 1998 (1998-07-16)	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

In rtional Application No
PCT/GB 99/02596

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9627675	A	12-09-1996	AU	4884396 A	23-09-1996
WO 9806737	A	19-02-1998	AU	4327797 A	06-03-1998
			CA	2232242 A	19-02-1998
			EP	0886644 A	30-12-1998
WO 9711094	A	27-03-1997	AT	184613 T	15-10-1999
			AU	4482996 A	09-04-1997
			CA	2232727 A	27-03-1997
			DE	69604298 D	21-10-1999
			EP	0851874 A	08-07-1998
US 5625048	A	29-04-1997	AU	702205 B	18-02-1999
			AU	4155096 A	21-08-1996
			CA	2205006 A	08-08-1996
			DE	29522103 U	30-09-1999
			EP	0804457 A	05-11-1997
			JP	10509881 T	29-09-1998
			WO	9623810 A	08-08-1996
			US	5777079 A	07-07-1998
WO 9830715	A	16-07-1998	AU	5090498 A	03-08-1998